**Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi**

# Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi

**Alessandro Desirò[1], Alessandra Salvioli[1], Eddy L. Ngonkeu[2], Stephen J. Mondo[3], Sara Epis[4], Antonella Faccio[5], Andres Kaech[6], Teresa E. Pawlowska[3] and Paola Bonfante[1]***

1. Department of Life Sciences and Systems Biology, University of Torino, Torino, Italy
2. Institute of Agronomic Research for Development (IRAD), Yaoundé, Cameroon
3. Department of Plant Pathology & Plant Microbe-Biology, Cornell University, Ithaca, New York, USA
4. Department of Veterinary Science and Public Health, University of Milano, Milano, Italy
5. Institute of Plant Protection, UOS Torino, CNR, Torino, Italy
6. Center for Microscopy and Image Analysis, University of Zurich, Switzerland

Correspondence: Paola Bonfante, Department of Life Sciences and Systems Biology, University of Torino, Viale P.A. Mattioli, 25, Torino 10125, Italy. E-mail: p.bonfante@ipp.cnr.it

***Abstract***

Arbuscular mycorrhizal fungi (AMF) are important members of the plant microbiome. They are obligate biotrophs that colonize the roots of most land plants and enhance host nutrient acquisition. Many AMF themselves harbor endobacteria in their hyphae and spores. Two types of endobacteria are known in Glomeromycota: rod-shaped

Gram-negative *Candidatus* Glomeribacter gigasporarum, *Ca*Gg, limited in distribution to members of the order Gigasporales, and coccoid *Mollicutes*-related endobacteria, *M*re, widely distributed across different lineages of AMF. The goal of the present study is to investigate the patterns of distribution and coexistence of the two endosymbionts, *Ca*Gg and *M*re, in spore samples of one host species, *Gigaspora margarita*. Based on previous observations, we hypothesized that some AMF could host populations of both endobacteria. To test this hypothesis we performed an extensive investigation of both endosymbionts in *G. margarita* spores sampled from Cameroonian soils as well as in the Japanese *G. margarita* MAFF520054 isolate using different approaches (molecular phylotyping, electron microscopy, fluorescence *in situ* hybridization, and quantitative real-time PCR). We found that a single AMF host can harbour both types of endobacteria, with *M*re population being more abundant, variable and prone to recombination than the *Ca*Gg one. Both endosymbionts seem to retain their genetic and lifestyle peculiarities regardless of whether they colonize the host alone or together. These findings show for the first time that fungi, as other eukaryotic hosts, support complex intracellular bacterial microbiomes, in which distinct types of endobacteria coexist in a single cell.

**Subject Category:** Microbial population and community ecology

**Keywords:** arbuscular mycorrhizal fungi; bacterial recombination, fungal microbiome; endobacteria; fluorescent *in situ* hybridization; phylogenetic analysis; quantitative real-time PCR

## Introduction

The discovery that the human body can be described as a complex ecosystem where human cells interact with trillions of bacteria and other microbes has represented a scientific revolution. The human microbiome, *i.e.* the microbial communities and the genetic information they contain, cooperate with the human genome to regulate crucial physiological processes ranging from digestion to obesity and immunity (Methé *et al*., 2012). Similarly, plants rely on microorganisms living both in their tissues and in the rhizosphere (Porras-Alfaro & Bayman, 2011; Berendsen *et al*., 2012, Lundberg *et al*., 2012; Bulgarelli *et al*., 2012). Due to their interdependence and mutual impact on each other's biology, plants and their microbiomes can be viewed as "super-organisms". To date, most of the work on plant-associated microbes focused almost exclusively on bacteria (Lundberg *et al*., 2012; Bulgarelli *et al*., 2012), even though eukaryotes such as fungi are also crucial components of the plant microbiome. They not only thrive in the rhizosphere, but also colonize plant tissues, exhibiting a range of lifestyles, including mutualism, parasitism and commensalism (Porras-Alfaro & Bayman, 2011).

Among plant-associated microbiota, arbuscular mycorrhizal fungi (AMF) are the most widespread: they belong to an ancient monophyletic phylum, the Glomeromycota (Schüßler *et al*., 2001), and play a key role in nutrient cycling and plant health due to their capacity for improving the mineral nutrition of plants (Smith & Read, 2008). AMF display many unusual biological features. In addition to their obligate biotrophy (Bonfante & Genre, 2010), many of them harbor endobacteria in their cytoplasm (Bonfante & Anca, 2009). Bacterial endosymbionts are widespread among animals (Wernegreen *et al*., 2012; McFall-Ngai *et al*., 2013), and in particular the ones living in insect tissues have been investigated in depth (Ferrari & Vavre,

2011). In contrast, examples of endobacteria living inside fungal cells are much more limited (Bianciotto *et al*., 2003; Partida-Martinez & Hertweck, 2005; Lackner *et al*., 2009; Naumann *et al*., 2010; Kai *et al*., 2012).

The endobacteria of Glomeromycota are the most thoroughly investigated bacterial endosymbionts of fungi, having been discovered in the early 1970s on the basis of electron microscope observations (Mosse, 1970). Two types of endosymbionts are known in AMF: (i) a rod-shaped, Gram-negative beta-proteobacterium (Bonfante *et al*., 1994), *Candidatus* Glomeribacter gigasporarum (*Ca*Gg), common in several species of the order Gigasporales (Bianciotto *et al*., 2003; Mondo *et al.,* 2012), and (ii) a coccoid bacterium displaying a homogeneous Gram-positive-like wall structure (MacDonald *et al*., 1982; Scannerini & Bonfante, 1991), which represents a currently undescribed taxon of *Mollicutes*-related endobacteria (*M*re) with a wide distribution across Glomeromycota (Naumann *et al*., 2010).

The *Ca*Gg genome sequence (Ghignone *et al*., 2012) revealed that Glomeribacter endobacteria are nutritionally dependent on the fungal host and have a possible role in providing the fungus with essential factors like vitamin B12 (Ghignone *et al*., 2012). Phenotypic consequences of *Ca*Gg removal from the host include important morphological changes as well as reduced proliferation of host presymbiotic hyphae. Yet, the host is not obligately dependent on the bacteria (Lumini *et al*., 2007; Mondo *et al*., 2012). These features suggest that Glomeribacter endobacteria are mutualistic associates of AMF (Lumini *et al*., 2007). Comparisons of host and symbiont phylogenies indicate that, while *Ca*Gg is a heritable endosymbiont (Bianciotto *et al*., 2004), it also engages in recombination and host switching, which play an important role in stabilizing this 400-million-year-old association (Mondo *et al*., 2012). In contrast, information on the coccoid *M*re is much more limited. Based

on the 16S rRNA gene sequences, this novel lineage is sister to a clade encompassing the Mycoplasmatales and Entomoplasmatales (Naumann *et al*., 2010). The *M*re have been detected in 17 out of 28 investigated AMF samples from culture collections, including members of Archaeosporales, Diversisporales, Gigasporales and Glomerales (Naumann *et al*., 2010), as well as in mycorrhizal thalli of liverworts (Desirò *et al*., 2013). In most of the AMF hosts and irrespectively of the AMF identity, these endobacteria displayed a conspicuous variability in their 16S rRNA gene sequence. Collectively, these observations indicate that *Ca*Gg is a stable associate of Gigasporales, while the lifestyle of the *M*re and the nature of their association with Glomeromycota are uncertain. Furthermore, the interaction between the two endosymbionts remains unclear, *i.e.* it is not known whether the presence of one endosymbiont in the host leads to the exclusion of the other one.

The goal of the present study is to investigate the patterns of distribution and coexistence of the two endosymbionts, *Ca*Gc and *M*re, in isolates of one host species, *Gigaspora margarita* W.N. Becker & I.R. Hall. Previous electron microscopy observations revealed that the strain of *G. margarita* MAFF520054 harboured a Gram-positive-like endobacterium (Kuga *et al*., 2008), while molecular analysis indicated the presence of *Ca*Gg (E. Lumini, personal communication, ref. seq. AM886455). Based on these observations, we hypothesize that some AMF could host populations of both endobacteria. To test this hypothesis we performed an extensive investigation of both endosymbionts in *G. margarita* spores sampled from Cameroonian soils as well as in *G. margarita* MAFF520054 from Japan using different approaches. We found that a single AMF host can harbour both types of endobacteria, with *M*re populations being more abundant, variable and prone to recombination than the *Ca*Gg ones. These findings show for the first time that fungi,

as other eukaryotic hosts, support complex microbiomes, in which distinct types of endobacteria coexist in a single cell.

## *Materials and methods*

All the details of the experimental procedures are available in the Supplementary Text S1.

### *Sampling and sample preparation*

Twelve soil samples were collected from three locations in Cameroon (Table 1). Trap cultures with *Sorghum* and *Vigna* were established using autoclaved sand mixed with the sampled soils. The Japanese isolate *G. margarita* MAFF520054 was provided by NIAS Genebank and propagated in pot cultures with *Trifolium*.

The spores were recovered from pot cultures by wet sieving (Gerdemann & Nicolson, 1963) and surface-sterilized (Lumini *et al*., 2007). The spore samples were morphologically identified as *Gigaspora margarita* following Bentivenga & Morton (1995).

### *DNA extraction, amplification, and clone library analysis*

DNA extractions were performed by crushing either individual spores or groups of five or ten spores according to Lumini *et al*., (2007). Three fragments of the fungal ribosomal gene cluster, namely 18S, ITS and 28S, were amplified.

The *Ca*Gg 16S rRNA gene was specifically amplified with the newly designed primers CaGgADf (5'-AGATTGAACGCTGGCGGCAT-3') and CaGgADr (5'-ATGCGTCCTACCGTGGCCATC-3'), while the *M*re 16S rRNA gene was amplified as described in Desirò *et al.* (2013).
Fungal and bacterial PCR amplicons were then cloned and transformed.

*Bioinformatic analyses*

Sequences were assembled and curated in Mega (Tamura *et al*., 2011), aligned with MAFFT (Katoh *et al*., 2002) or MUSCLE (Edgar, 2004), and then examined for chimerism. Sequence similarity/divergence was evaluated using MOTHUR (Schloss *et al.,* 2009). Nucleotide diversity ($\pi$) was calculated in DNAsp v. 5.10.01 (Librado & Rozas, 2009). The *Ca*Gg and *M*re 16S rRNA gene sequences were grouped into operational taxonomic units (OTUs) at the cutoff of 0.03 genetic distance value using MOTHUR. Phylogenetic analyses were conducted using one representative sequence for each OTU. The Genetic Algorithm for Recombination Detection, GARD (Kosakovsky Pond *et al*., 2006), was used to identify recombination breakpoints in 16S rRNA genes of *Ca*Gg and *M*re.. Alignments and trees are available in TreeBASE (submissions XXXX) (Piel *et al*., 2002). Representative DNA sequences are in GenBank (XXXX).

*Ultrastructural analyses*

Single *G. margarita* spores from CM23 and CM 47 samples were processed by using high-pressure-freezing followed by freeze-substitution. Single spores floating in water were transferred in the cavity of an aluminium carrier with a

pipette. Excess of water was drawn off with filter paper and the space was filled with 1-Hexadecene. The sandwich was completed with a flat specimen carrier and frozen in a HPM 100 high-pressure freezing machine (Leica Microsystems, Wetzlar, Germany) (McDonald *et al*., 2010). Samples were then freeze-substituted, resin embedded, and processed for transmission electron microscopy.

*FISH experiments and Confocal Microscopy*

Sterilized spores of the samples CM23, CM47, CM50, CM52 and *G. margarita* BEG34 were fixed as described in Naumann *et al*., (2010). The *M*re specific probe BLOsADf2 (Desirò *et al*., 2013), together with a newly designed specific *Ca*Gg 16S rRNA probe (CaGgADf1 5'-CTATCCCCCTCTACAGGAYAC-3'), were used to label the endobacteria. In addition, the eubacterial probe EUB338 (Amann *et al*., 1990) and the *Buchnera*-specific probe ApisP2a (Koga et al., 2003) were used. Spores were observed using a Leica TCS-SP2 confocal microscope (Leica Microsystems).

*Quantification of the bacterial populations*

The sample CM23 (containing both *M*re and *Ca*Gg) was selected for the relative quantification of the two bacterial populations by real-time qPCR. Briefly, the 16S rDNA gene sequences obtained for both *Ca*Gg and *M*re were used to design two distinct qPCR primer pairs. Template plasmids containing the target DNA sequences were constructed to generate a standard curve as an external standard. The number of target DNA sequences present in each

PCR mixture was calculated by comparing the crossing points of the samples with those of the standards.

## ***Results***

### *Identity of AMF*

To confirm the morphological identification of AMF originating from Cameroon and Japan as *Gigaspora margarita*, we analysed their 18S, 28S and ITS rRNA gene regions. These analyses revealed that all the fungi could be identified as *G. margarita* (Figures 1 and S3). As expected, the 18S rRNA gene analysis led to an unresolved, polytomic phylogeny (not shown), while a better resolution was provided by the 28S rRNA gene (Figure 1) and the ITS region (Figure S3).

### *Identity of endobacteria*

Bacterial 16S rRNA gene sequences were PCR-amplified from single AMF spores using primers specific for *Ca*Gg and *M*re (Naumann ***et al***., 2010) to detect endosymbiont presence. Most samples harboured both types of endobacteria with the exception of the *G. margarita* samples CM3 and CM52, which contained only *M*re (Table 1). The absence of *Ca*Gg in the samples CM3 and CM52 was confirmed by real-time qPCR (data not shown), which can detect up to ten bacterial cells (Salvioli ***et al***., 2008).

In order to faithfully describe the microbiome contained inside the AMF spores and to capture all of the bacterial biodiversity, a more extensive analysis was performed on pools of ten spores from four Cameroonian samples (CM21, CM23, CM47, CM50) and from the Japanese isolate.

The RFLP analyses of *Ca*Gg 16S rRNA gene sequences revealed a single RFLP profile for each 10-spores sample, suggesting a limited intrasample variability, which was further confirmed by sequence analyses. The obtained sequences were grouped into OTUs at 97% of sequence similarity and, as expected, a single OTU for each sample was obtained (Table 2). Phylogenetic analyses of *Ca*Gg sequences retrieved from spore samples showed that they clustered with other *Ca*Gg sequences available in GenBank (Figure 2).

Sequencing of the *M*re 16S rRNA gene clones generated a total of 118 sequences (Table 3). To eliminate potential PCR artefacts expected in amplifications from complex templates such as *M*re populations (Naumann *et al*., 2010), the obtained sequences were submitted to a rigorous chimera screen, which reduced the total amount to 52 sequences (Table 3). They were grouped into OTUs at 97% sequence similarity (Table 3). Most of the sequences (48 out of 52) showed sequence similarity values lower than 97% when compared to the *M*re sequences obtained from GenBank, suggesting the presence of novel phylotypes (Table 3).

Despite the high variability, all retrieved *M*re sequences clustered together with those obtained in previous studies (Naumann *et al*., 2010; Desirò *et al*., 2013) (Figure 3). Moreover, because the resulting phylogenies presented here are better supported and resolved than those constructed in previous works (Naumann *et al*., 2010; Desirò *et al*., 2013), we conclude that there are at least two distinct and well supported *M*re clades, identified as *M*re group A and *M*re group B (Figure 3), and that the level of sequence divergence among sequences clustering in the same *M*re group reached up to

15 and 16% in *M*re group A and B, respectively. Overall, in all the samples, with the only exception of CM50, *Ca*Gg showed a high level of intra-host sequence similarity, whereas *M*re revealed high levels of intra-host sequence diversity.

*Recombination detection*

To explore the underlying causes of differences in sequence evolution patterns between *Ca*Gg and *M*re, we used GARD (Kosakovsky Pond *et al*., 2006) to look for evidence of recombination in 16S rRNA genes of the two endosymbionts associated with AMF from Cameroon and Japan. No evidence of recombination was detected in the *Ca*Gg sequences. In contrast, in the *M*re dataset, we found that the $AIC_C$ score of 8529.9 for the best-fitting model allowing for different topologies of the alignment segments defined by recombination breakpoints was lower than the $AIC_C$ score of 8819.4 for the model that assumed the same topology for all segments, indicating that a multiple tree model is preferable over a single tree model. Using the KH test, one breakpoint at the alignment position 479 was identified as resulting in significant topological incongruence between segments ($P < 0.001$, Figure S4).

*Localization of the two bacterial morphotypes in AMF cells: high pressure/freeze-substitution and transmission electron microscopy*

We used electron microscopy to confirm the cytoplasmic location of both types of endobacteria. To ensure proper preservation of endosymbiont cells and fungal organelles, which could be jeopardized by the very thick fungal cell

wall (12-16 µm, Lumini *et al*., 2007), we used high pressure and freeze-substitution specimen preparation. On the basis of the previous molecular analysis, two isolates of *G. margarita* (CM23 and CM47) were selected for this experiment. When inspected under the electron microscope, CM23 and CM47 presented both the rod-shaped and coccoid bacteria in the same area of their cytoplasm (Figure 4). The rod-shaped *Ca*Gg were 330-550 x 960-1050 nm in size, with a layered, Gram-negative type cell wall (Figure 4A, B) and were located inside a vacuole-like organelle (Figure 4A), consistent with reports from earlier studies (Bianciotto *et al*., 1996, 2003). The vacuole revealed an electron dense matrix, which was identified as of protein origin (Bonfante *et al*., 1994) (Figure 4A). In other cases, the matrix was reduced in size and the bacterium was more closely surrounded by the membrane of fungal origin (Figure 4B). In contrast, the coccoid *M*res were directly embedded in the fungal cytoplasm (Figure 4A, C). They were consistently smaller, 300-600 nm in size, with a homogeneous, Gram-positive-like cell wall (Figure 4C).

*Localization of the two endosymbionts in AMF spores: FISH*

To further validate our molecular and morphological observations of the *Ca*Gg and *M*re coexistence in *G. margarita,* we performed fluorescence *in situ* hybridization (FISH) experiments in samples CM23, CM47, CM50, and CM52. *G. margarita* BEG34 was used as negative control, since *M*re have never been found in this isolate (Naumann *et al*., 2010). We used two probes: CaGgADf1, which was designed to specifically detect *Ca*Gg, and BLOsADf2 (Desirò *et al*., 2013), which targeted entire *M*re variability contained in our spore samples. In agreement with PCR results, we did not observe any *Ca*Gg

signal in CM52, where *Ca*Gg have never been detected by PCR-amplification of 16S rRNA gene. Similarly, we did not observe any *M*re signals in BEG34. On the contrary, the two specific probes produced simultaneous FISH signals in the spores where the presence of both bacterial types was expected (Figures 5 and 6). The number of fluorescent signals suggested a more abundant presence of *M*re than *Ca*Gg in the spores with both types of bacteria. The fluorescent signals were located in the fungal cytoplasm and never on the spore surface. Importantly, the fluorescent signal of the probes BLOsADf2 (Desirò *et al*., 2013) and CaGgADf1 were always co-localized with the fluorescence given by the general bacterial probe EUB338 (Amann *et al*., 1990) (Fig. 5). No fluorescent signal was detected with the negative control probe ApisP2a (Koga *et al*., 2005) (Fig. 6E). Pre-treatment with RNase, as well as control hybridization with nonsense probes, did not provide any FISH signal. A weak autofluorescence of the fungal cytoplasm, probably deriving to the use of aldehydic fixatives, was visible in all spore samples. Hence, FISH experiments, validating the PCR results, confirmed the simultaneous presence of *M*re and *Ca*Gg in some *G. margarita* samples.

*M*re *and CaGg abundance in AMF cells: real-time qPCR*

To further examine differences in *M*re and *Ca*Gg abundance suggested by FISH experiments, we used real-time quantitative PCR (qPCR) to quantify the bacterial populations present in the *G. margarita* sample CM23 that was previously shown to contain both *M*re and *Ca*Gg endobacteria. The 16S rRNA gene was used as a target gene, but while in the *Ca*Gg genome the 16S rRNA gene is present in a single copy (Ghignone *et al*., 2012), in *M*re one or

at most two rRNA gene copies are expected based on the comparison with the closest microbes already sequenced (Fraser *et al*., 1995; Glass *et al*., 2000; Jaffe *et al*., 2004; Minion *et al*., 2004; Vasconcelos *et al*., 2005; Bai *et al*., 2006).

The accuracy of qPCR primers of *Ca*Gg and *M*re was confirmed by assessing the melting profile generated by each primer pair (Figure S2). Subsequently, we quantified the relative abundance of the two bacterial endosymbionts on the basis of the 16S rRNA gene sequences. In *G. margarita* CM23, we found that *M*re were always more abundant than *Ca*Gg, and the bacterial ratio was maintained fairly constant irrespective of the size of the batches considered (*i.e.* one, five or ten spores) (Table 4).

The qPCR analysis of the bacterial 16S rRNA gene sequences revealed that *M*re are 5.17 - 6.12 times more abundant than *Ca*Gg in the *G. margarita* CM23 spores, assuming that a single 16S rRNA gene is present in the *M*re genome. This value should be reduced to 2.59 - 3.06 times if two copies of the 16S rRNA are present in *M*re genomes instead (Table 4). This finding is consistent with our FISH observations, which suggested that *M*re were more abundant than *Ca*Gg in *G. margarita* spores.

## *Discussion*

A combination of morphological, molecular, and phylogenetic analyses demonstrates that *Gigaspora margarita* spores host a complex microbiome consisting of rod-shaped and coccoid bacteria. The two bacterial groups are very distinct not only in their phylogenetic placement, *i.e.* C*andidatus* Glomeribacter gigasporarum is closely related to Burkholderiaceae, while the

coccoid endobacteria are related to the Gram-positive *Mollicutes*, but also in their genetic features.

*Sharing the same host and revealing intra-host diversity*

Notwithstanding the endobacteria share the same fungal host, a relevant difference in genetic diversity patterns between them was revealed. While *Ca*Gg shows a high level of intra-host sequence similarity, the *M*re are characterized by high levels of intra-host sequence diversity. One of the underlying causes of differences in sequence evolution patterns between *Ca*Gg and *M*re may be differences in their lifestyle. For example, in *M*re, we found evidence of recombination, which was not apparent in *Ca*Gg. This finding was supported by some genomic features of *Ca*Gg genome: notwithstanding its high repetitive DNA (15%), CaGg contains a low number of active insertion sequences, which are considered important determinants for recombination (Ghignone *et al*., 2012). Indeed, a recent study of *Ca*Gg, using a set of four marker genes, revealed that recombination is not entirely absent from the *Ca*Gg evolutionary history and, together with host switching, may play an important role in evolutionary stability of *Ca*Gg association with Glomeromycota (Mondo *et al*., 2012). Detecting evidence of recombination in a single gene of *M*re sampled in the present study may suggest that *M*re engage in more frequent recombination than *Ca*Gg. Interestingly, cryptic prophage remnants have been detected in the genome of the *M*re-related phytoplasma, leading to the suggestions that these genetic elements may have played important roles in generating phytoplasma genetic diversity (Wei *et al*., 2008).

*Phylogenetic divergence patterns of the co-existing endobacteria*

The extensive phylogenetic analysis performed on the endobacteria thriving in the cytoplasm of five spore samples and their comparison with data from previous investigations (Bianciotto *et al*., 1996, 2000, 2003; Mondo *et al*., 2012) confirmed that the 16S rRNA gene sequences of *Ca*Gg were relatively conserved, irrespectively of the geographic origin of the fungal host. However, our careful analyses showed that the sequence similarity between *Ca*Gg from *G. margarita* MAFF520054 isolate and the already sequenced *Ca*Gg from *G. margarita* BEG34 was below the critical level of 97%. In fact, although this distinction is controversial (Rossello-Mora, 2003), it is generally accepted that sequences with similarity greater than 97% are typically assigned to the same species and those with similarity greater than 95% to the same genus (Stackebrandt & Goebel, 1994; Everett *et al*., 1999; Gevers *et al*., 2005). Consequently, further work is needed to resolve whether *Ca*Gg from *G. margarita* MAFF520054 and *G. margarita* BEG34, which show sequence similarity lower than 97% and a different location inside the *Ca*Gg phylogenetic tree, represent distinct taxa.

In contrast to *Ca*Gg and despite the stringent removal of chimeric sequences, the 16S rRNA gene sequences of *Mollicutes*-related endobacteria turned out to be highly variable inside at least four out of five spore samples. Moreover, in only 8% of the sequences generated in this study (4 out of 52), the similarity with sequences from GenBank was above 97%; the remaining 92% of the sequences showed sequence similarity lower than 97%. Despite such high sequence dissimilarity levels, all *M*re sequences obtained in this

study clustered together with the ones previously retrieved from Glomeromycota spore collection and liverworts-associated AMF. It is additionally possible that the stringent chimera removal excluded some non-chimeric sequences. However, this allowed us to enhance our phylogenetic resolution beyond what was presented in previous studies (Naumann *et al*., 2012; Desirò *et al*., 2013). As a result, we could recognize at least two distinct well supported *M*re clades, here identified as *M*re group A and *M*re group B. However, due to high level of sequence divergence between *M*re sequences clustering in the same *M*re group, we hypothesize that these newly described groups can mask other still hidden clades.

*Morphological aspects of endobacteria are not affected by their co-occurrence*

Our present study is the first one to describe in a single fungal host the coexistence of two distinct bacterial endosymbionts. Until now, these two symbionts have been studied in isolation from each other. We found that the morphological characteristics of the two coexisting bacterial endosymbionts did not differ from those described previously in the samples where only one bacterial symbiont was present. For example, even when sharing the same cell volume, *Ca*Gg remained enclosed in a vacuole-like structure, while *M*re were embedded directly in the cytoplasm.

Interestingly, the spore samples that we investigated showed different patterns of intersymbiont dynamics. For example, in the sample CM50, only one *M*re phylotype revealed high values of sequence similarity and consequently a limited number of *M*re single OTUs was detected together with

the homogenous *Ca*Gg population. In contrast, in the remaining samples, *M*re showed higher levels of nucleotide diversity and sequence divergence. It would be useful to explore which of these two scenarios is more recent and which is more evolutionarily stable.

Irrespective of the dynamic levels of *M*re sequence similarity in different samples, FISH and molecular quantitative analysis revealed that *M*re were unambiguously more abundant than *Ca*Gg. The stronger presence of the *M*re together with their high variability, may indicate that they are more aggressive colonizers of AMF. On the basis of their 16S rDNA phylogeny, *M*re have been described as related to *Mollicutes* (Naumann *et al*., 2010), a bacterial group that clusters with microbes (*i.e. Mycoplasma*) thriving inside many eukaryotic hosts and manipulating host development thanks to the release of effector proteins (Sugio *et al*., 2011). Due to their capacity to interact with many AM host genotypes, we hypothesize that *M*re have been one of the factors shaping AMF evolution and/or their ecological success.

*Similarities between endosymbionts of insects and AMF*

The wealth of natural history and molecular evolution data available for heritable endosymbionts of insects make them into an excellent model for understanding symbiotic associations that involve vertically transmitted endobacteria. In addition to essential endosymbionts, insects can support complex communities of bacteria that include non-essential endosymbionts as well as reproductive manipulators (Moran *et al*., 2008). Essential endosymbionts show strict vertical transmission and functional complementation with their hosts resulting from millions of years of reciprocal

selection (McCutcheon & Moran, 2010). The genomes of essential endosymbionts are usually highly reduced (McCutcheon & Moran, 2010; McFall-Ngai *et al*., 2013). In this context, *Buchnera aphidicola* is a paradigm for primary endosymbionts. *Buchnera*'s association with aphids is ancient, being approximately 200 million years old and revolves around the endosymbiont's capacity to synthesize essential amino acids for its host (van Ham *et al*., 2003). Due to their pleiotropic effects on their hosts, the situation is not so clear-cut for the non-essential (secondary) endosymbionts, since their transmission may be both vertical and horizontal and the ratio between cost and benefits strictly depends on environmental conditions (Ferrari & Vavre, 2011). ~~The effects of secondary symbionts on their host are pleiotropic, but one of the best understood is their action in the so-called tritrophic interactions. For example, the secondary symbiont *Hamiltonella defensa* confers on its aphid host an increased level of resistance towards the parasitoid wasp when compared to the uninfected aphids of the same genotype (Oliver *et al.*, 2003).~~

~~Reproductive parasites manipulate their insect host reproduction in favour of their own transmission (Engelstädter & Hurst, 2009). Examples of reproducitve manipulators include *Wolbachia* and *Rickettsia* (Engelstädter & Hurst, 2009) as well as *Spiroplasma* (Anbutsu & Fukatsu, 2011). Their lifestyles can be highly dynamic showing fast transitions between parasitism and mutualism (Weeks *et al.*, 2007).~~

Given our observations that a single cell (a spore) of a fungus can host endosymbionts with distinct characteristics, it is worth considering whether the

biological features of these fungal endobacteria are comparable to those of endosymbionts of insects.

In the case of *Ca*Gg, one of its hosts, *Gigaspora margarita*, can survive and multiply in the absence of the endobacterium (Lumini *et al*., 2007), and there are natural *Ca*Gg-free isolates of Gigasporales (Mondo *et al*., 2012), demonstrating that this symbiosis is facultative for the host. However, the fungal fitness can be strongly reduced by removal of the endobacteria (Lumini *et al*., 2007; P. Bonfante and M. Novero 2013, unpublished data). In addition, by using codiverging partner pairs, Mondo *et al*. (2012) demonstrated that this fungal/bacterial association is ancient (at least 400 million years old) and evolutionarily stable. Analysis of the 1.72 Mb *Ca*Gg genome (Ghignone et *al.*, 2012) revealed that it is reduced when compared with the free-living related *Burkholderia* species, and that the metabolic profile of *Ca*Gg unambiguously clusters with insect endobacteria, including essential endosymbionts like *Buchnera* and *Wigglesworthia* (Moran *et al*., 2008). These data suggest that *Ca*Gg has undergone functional convergent evolution with phylogenetically distant endobacteria. However, genome annotation also shows functional similarities with the secondary non-essential symbionts (for example *H. defensa*). On the basis of these considerations, we concluded that *Ca*Gg is an obligate intracellular symbiont, characterized by a genetic mosaic where determinants for different nutritional strategies are integrated in a reduced genome (Ghignone *et al*., 2012). Collectively, its life history features (*i.e.* a strict vertical transmission) as well as molecular evolution and genomic features seem to share patterns from both essential and non-essential endosymbionts of insects.

While the knowledge of the *M*re biology is too limited to advance any hypothesis concerning their impact on the host biology, *M*re relatedness to *Mycoplasma* and *Phytoplasma*, which are widespread parasites of animals and plants, might explain the colonization capacities of *M*re, irrespectively of their role in the fungal hosts. On the other hand, it cannot be excluded that they are beneficial associates of fungi, akin to *Spiroplasma* endosymbionts that protect their insect hosts from the parasitoid pressure (Xie *et al*., 2010). Consequently, taken in consideration the limited available empirical evidence, we conclude that classifying *M*re into categories established for bacterial associates of insects is not yet possible.

*Are endobacteria favoured by coenocytic hyphae?*

In the rapidly evolving taxonomic classification of Glomeromycota (Redecker *et al*., 2013), the taxon named Gigasporales (Oehl *et al*., 2011; da Silva *et al*., 2013) identifies a group of AMF with distinct features of spore morphology (size, wall layering, bulbous base, germination shield) and host root colonization patterns (lack of intraradical vesicles and formation of auxiliary cells). In addition, this lineage of Glomeromycota turns out to be a preferential niche for endobacteria. Our present results confirm previous analyses (Bianciotto *et al*., 1996, 2000, 2003; Mondo *et al*., 2012) that demonstrated a strict association of *Ca*Gg with the Gigasporales. In contrast, the *M*re are widespread; they have been found in both basal and more recently evolved Glomeromycota taxa (Naumann *et al*., 2010). This differential distribution pattern is one of the key distinctions between the two groups of endosymbionts.

Our present results clearly demonstrate that *Gigaspora margarita* can harbour both endosymbionts, *Ca*Gg and *M*re, and this is probably true also for other Gigasporales taxa (A. Desirò and G. A. da Silva 2013, personal communication). The underlying mechanisms responsible for the propensity of Gigasporales to host endobacteria are unknown. However, the genome sequence of the *Ca*Gg (Ghignone *et al*., 2012) shows that this bacterium is metabolically dependent on its fungal host. Perhaps only Gigasporales with their relatively large spores, which are rich in reserves of glycogen, fats, and proteins (Bonfante *et al*., 1994), can support the energetic cost of complex bacterial communities, which thrive inside a protected niche.

There is, however, increasing evidence that *Mortierella* species (Mucoromycotina) host endobacteria that are related to *Ca*Gg (Sato et *al*., 2010; Kai *et al*., 2012; Bonito *et al*., 2013). These data open a novel interesting scenario: fungal endobacteria might prefer coenocitic hyphae. The absence of transverse septa may facilitate bacterial movement across the fungal mycelium, as observed in the *Burkholderia rhizoxinica* endosymbiont of *Rhizopus microsporus* (Partida-Martinez & Hertweck, 2005). In addition, these data support a link between Glomeromycota and Mucoromycotina, which both belong to the group of the basal fungi: this has already been suggested by the similarities in their mitochondrial genomes (Lee & Young, 2009; Pelin *et al.*, 2012) and by the assembled genome of Rhizophagus irregularis (M.Martin, Ton Bisseling..personal communication). This pattern of endosymbiont distribution across lineages of closely related fungal hosts raises questions about the role of symbiosis in the evolution and diversification of these fungal taxa and their associated endobacteria.

## Conclusion

Our investigation has revealed for the first time that a single spore of an AMF can harbour multiple bacterial endosymbionts that represent phylogenetically diverse groups and show distinct patterns of sequence evolution. Both endosymbionts seem to retain their genetic and lifestyle peculiarities regardless of whether they colonize the host alone or together. *M*re population consistently appears to be more abundant, variable and prone to recombination events than the *Ca*Gg one, suggesting that the same niche (the fungal spore) exerts a different selection pressure on its dwellers.

Our findings showing that a single fungal cell can harbour a complex microbiome, raise novel questions concerning molecular, cellular and metabolic interactions resulting from such complex inter-domain relationships.

## Acknowledgement

The Authors thank Andrea Genre (DiBIOs, UNITO) for his invaluable assistance in confocal microscopy and image acquisition; Gladstone Alves da Silva (Mycology Department , Pernanbuco University) for the morphological identification of the spores; Yukari Kuga (Agriculture Faculty, Shinshu University) for her suggestions in the endobacterial ultrastructure; Stefano Ghignone and Olivier Friard (IPP-CNR and DiBIOs, UNITO) for their assistance in bioinformatics; the Genebank project of NIAS (Japan) for providing the MAFF520054 isolate. Research was funded by the Ateneo Project (ex 60%) to PB.

Supplementary information is available at ISMEJ's website.

## References


Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56:** 1919-1925.

Anbutsu H, Fukatsu T. (2011). Spiroplasma as a model insect endosymbiont. *Environ Microbiol* **3:** 144-153.

Bai X, Zhang J, Ewing A, Miller SA, Radek AJ, Shevchenko DV *et al*. (2006). Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. *J Bacteriol* **188:** 3682-3696.

Bentivenga SP, Morton JB. (1995). A monograph of the genus *Gigaspora*, incorporating developmental patterns of morphological characters. *Mycologia* **87:** 719-731.

Berendsen RL, Pieterse CM, Bakker PA. (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478-486.

Bianciotto V, Genre A, Jargeat P, Lumini E, Becard G, Bonfante P. (2004). Vertical transmission of endobacteria in the arbuscular mycorrhizal fungus *Gigaspora margarita* through generation of vegetative spores. *Appl Environ Microbiol* **70:** 3600-3608.

Bianciotto V, Lumini E, Bonfante P, Vandamme P. (2003). '*Candidatus* Glomeribacter gigasporarum' gen. nov., sp. nov., an endosymbiont of arbuscular mycorrhizal fungi. *Int Syst Evol Micr* **53:** 121-124.

Bianciotto V, Lumini E, Lanfranco L, Minerdi D, Bonfante P, Perotto S. (2000). Detection and identification of bacterial endosymbionts in arbuscular

mycorrhizal fungi belonging to the family Gigasporaceae. *Appl Environ Microbiol* **66:** 4503-4509.

Bianciotto V, Bandi C, Minerdi D, Sironi M, Tichy HV, Bonfante P. (1996). An obligately endosymbiotic mycorrhizal fungus itself harbors obligately intracellular bacteria. *Appl Environ Microbiol* **62:** 3005-3010.

Bonfante P, Genre A. (2010). Mechanisms underlying beneficial plant-fungus interactions in mycorrhizal symbiosis. *Nat Commun* **1:** 48-58.

Bonfante P, Anca IA. (2009). Plants, mycorrhizal fungi, and bacteria: a network of interactions. *Annu Rev Microbiol* **63:** 363-383.

Bonfante P, Balestrini R, Mendgen K. (1994). Storage and secretion processes in the spore of *Gigaspora margarita* Becker & Hall as revealed by high-pressure freezing and freeze substitution. *New Phytol* **128:** 93-101.

Bonito G, Gryganskyi A, Schadt C, Pelletier D, Schaefer A, Tuskan G *et al.* (2013). Genomic analysis of *Mortierella elongata* and its endosymbiotic bacterium. 27$^{TH}$ fungal genetics conference, Asilomar.

Bulgarelli D, Rott M, Schaleppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F *et al*. (2012). Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* **488:** 91-95.

da Silva GA, Maia LC, Oehl F. (2013). Phylogenetics systematics of the Gigasporales. *Mycotaxon* **122:** 207-220.

Desirò A, Naumann M, Epis S, Novero M, Bandi C, Genre A, Bonfante P. (2013). *Mollicutes*-related endobacteria thrive inside liverwort-associated arbuscular mycorrhizal fungi. *Environ Microbiol* **15:** 822-836.

Edgar RC. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

Engelstädter J, Hurst GDD. (2009). The ecology and evolution of microbes that manipulate host reproduction. *Annu Rev Ecol Syst* **40:** 127-149.

Everett KDE, Bush R M, Andersen AA. (1999). Emended description of the order Chlamydiales, proposal of Parachlamydiaceae fam. nov. and Simkaniaceae fam. nov., each containing one monotypic genus, revised taxonomy of the family Chlamydiaceae, including a new genus and five new species, and standards for the identification of organisms. *Int J Syst Bacteriol* **49:** 415-440.

Ferrari J, Vavre F. (2011). Bacterial symbionts in insects or the story of communities affecting communities. *Philos Trans R Soc B* **12:** 1389-1400.

Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD *et al.* (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science* **270:** 397-403.

Gerdemann JW, Nicolson TH. (1963). Spores of mycorrhizal *Endogone* species extracted from soil by wet sieving and decanting. *Trans Br Mycol Soc* **46:** 235-244.

Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ *et al*. (2005). Re-evaluating prokaryotic species. *Nat Rev Microbiol* **3:** 733-739.

Ghignone S, Salvioli A, Anca I, Lumini E, Ortu G, Petiti L *et al*. (2012). The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. *ISME J* **6:** 136-45.

Glass JI, Lefkowitz EJ, Glass JS, Heiner CR, Chen EY, Cassell GH. (2000). The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. *Nature* **407:** 757-762.

Jaffe JD, Stange-Thomann N, Smith C, DeCaprio D, Fisher S, Butler J *et al*. (2004). The complete genome and proteome of *Mycoplasma mobile*. *Genome Res* **14:** 1447-1461.

Jargeat P, Cosseau C, Ola'h B, Jauneau A, Bonfante P, Batut J *et al*. (2004). Isolation, free-living capacities, and genome structure of *Candidatus* glomeribacter gigasporarum, the endocellular bacterium of the mycorrhizal fungus *Gigaspora margarita*. *J Bacteriol* **186:** 6876-6884.

Kai K, Furuyabu K, Tani A. (2012). Production of the quorum-sensing molecules *N*-acylhomoserine lactones by endobacteria associated with *Mortierella alpina* A-178. *Chembiochem* **13:** 1776-1784.

Katoh K, Misawa K, Kuma K, Miyata T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucl Acids Res* **30:** 3059-3066.

Koga R, Tsuchida T, Fukatsu T. (2003). Changing partners in an obligate symbiosis: a facultative endosymbiont can compense for loss of the essential endosymbiont *Bunchnera* in an aphid. *Proc R Soc B* **270:** 2543-2550.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* **23:** 1891-1901

Kuga Y, Saito K, Nayuki K, Peterson RL, Saito M. (2008). Ultrastructure of rapidly frozen and freeze-substituted germ tubes of an arbuscular mycorrhizal fungus and localization of polyphosphate. *New Phytol* **178:** 189-200.

Lackner G, Möbius N, Scherlach K, Partida-Martinez LP, Winkler R, Schmitt I, Hertweck C. (2009). Global distribution and evolution of a toxinogenic *Burkholderia-Rhizopus* symbiosis. *Appl Environ Microbiol* **75:** 2982-2986.

Lee J, Young JP. (2009). The mitochondrial genome sequence of the arbuscular mycorrhizal fungus *Glomus intraradices* isolate 494 and implications for the phylogenetic placement of *Glomus*. *New Phytol* **183:** 200-211.

Librado P, Rozas J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451-1452.

Lumini E, Bianciotto V, Jargeat P, Novero M, Salvioli A, Faccio A *et al.* (2007). Presymbiotic growth and sporal morphology are affected in the arbuscular mycorrhizal fungus *Gigaspora margarita* cured of its endobacteria. *Cell Microbiol* **9:** 1716-1729.

Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S *et al*. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488:** 86-90.

MacDonald RM, Chandler M, Mosse B. (1982). The occurrence of bacterium-like organelles in vesiculararbuscular mycorrhizal fungi. *New Phytol* **90:** 659-663.

McCutcheon JP, Moran NA. (2010). Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. *Genome Biol Evol* **2:** 708-718.

McDonald K, Schwarz H, Müller-Reichert T, Webb R, Buser C, Morphew M. (2010). “Tips and tricks” for high-pressure freezing of model systems. *Methods Cell Biol* **96:** 671-693.

McFall-Ngai M, Hadfield MG, Bosch TC, Carey HV, Domazet-Loso T, Douglas AE *et al.* (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci USA* **110:** 3229-3236.

Methé BA, Nelson KE, Pop M, Creasy HH, Giglio MG, Huttenhower C *et al.* (2012). A framework for human microbiome research. *Nature* **486**: 215-221.

Minion FC, Lefkowitz EJ, Madsen ML, Cleary BJ, Swartzell SM, Mahairas GG. (2004). The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swinemycoplasmosis. *J Bacteriol* **186:** 7123-7133.

Mondo SJ, Toomer KH, Morton JB, Lekberg Y, Pawlowska TE. (2012). Evolutionary stability in a 400-million-year-old heritable facultative mutualism. *Evolution* **66:** 2564-2574.

Moran NA, McCutcheon JP, Nakabachi A. (2008). Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet* **42:** 165-190.

Mosse B. (1970). Honey-coloured, sessile *Endogone* spores. *Arch Mikrobiol* **74:** 146–159.

Naumann M, Schüßler A, Bonfante P. (2010). The obligate endobacteria of arbuscular mycorrhizal fungi are ancient heritable components related to the *Mollicutes. ISME J* **4:** 862-871.

Oehl F, da Silva GA, Goto BT, Maia LC, Sieverding E. (2011). Glomeromycota: two new classes and a new order. *Micotaxon* **116:** 365–379.

Oliver KM, Russell JA, Moran NA, Hunter MS. (2003). Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc Natl Acad Sci USA* **100:** 1803-1807.

Partida-Martinez LP, Hertweck C. (2005). Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* **437:** 884-888.

Pelin A, Pombert JF, Salvioli A, Bonen L, Bonfante P, Corradi N. (2012). The mitochondrial genome of the arbuscular mycorrhizal fungus *Gigaspora*

*margarita* reveals two unsuspected trans-splicing events of group I introns. *New Phytol* **194:** 836-845.

Piel WH, Donoghue MJ, Sanderson MJ. (2002). TreeBASE: a database of phylogenetic knowledge. In: Shimura J, Wilson KL, Gordon D (eds). *To the interoperable "Catalog of Life" with partners species 2000 Asia Oceania*. National Institute for Environmental Studies: Tsukuba, pp 41-47.

Porras-Alfaro A, Bayman P. (2011). Hidden fungi, emergent properties: endophytes and microbiomes. *Annu Rev Phytopathol* **49:** 291-315.

Redecker D, Schüßler A, Stockinger H, Stürmer SL, Morton JB, Walker C. An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). *Mycorrhiza* 2013; e-pub ahead of print 5 April 2013, doi: 10.1007/s00572-013-0486-y.

Rossello-Mora R. (2003). Opinion: the species problem, can we achieve a universal concept? *Syst Appl Microbiol* **26:** 323-326.

Salvioli A, Lumini E, Anca IA, Bianciotto V, Bonfante P. (2008). Simultaneous detection and quantification of the unculturable microbe *Candidatus* Glomeribacter gigasporarum inside its fungal host *Gigaspora margarita*. *New Phytol* **180:** 248-257.

Sato Y, Narisawa K, Tsuruta K, Umezu M, Nishizawa T, Tanaka K *et al.* (2010). Detection of betaproteobacteria inside the mycelium of the fungus *Mortierella elongata*. *Microbes Env* **25:** 321-324.

Scannerini S, Bonfante P. (1991). Bacteria and bacteria-like objects in endomycorrhizal fungi. In: Margulis L, Fester R (eds). *Symbiosis as a source of evolutionary innovation: speciation and morphogenesis*. MIT Press: Cambridge, pp 273-287.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al*. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75:** 7537-7541.

Schüßler A, Schwarzott D, Walker C. (2001). A new fungal phylum, the Glomeromycota: phylogeny and evolution. *Mycol Res* **105:** 1413-1421.

Smith VSE, Read DJ. (2008). *Mycorrhizal Symbiosis 3rd edn*. Academic Press: San Diego.

Stackebrandt E, Goebel BM. (1994). A place for DNA-DNA reassociation and 16S rRNA sequence-analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44:** 846-849.

Sugio A, Kingdom HN, MacLean AM, Grieve VM, Hogenhout SA. (2011). Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. *Proc Natl Acad Sci USA* **108:** 1254-1263.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28:** 2731-2739.

van Ham RC, Kamerbeek J, Palacios C, Rausell C, Abascal F, Bastolla U *et al.* (2002). Reductive genome evolution in *Buchnera aphidicola*. *Proc Natl Acad Sci USA* **100:** 581-586.

Vasconcelos AT, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM *et al.* (2005). Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. *J Bacteriol* **187:** 5568-5577.

Weeks AR, Turelli M, Harcombe WR, Reynolds KT, Hoffmann AA. (2007). From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*. *PLoS Biol* **5:** 997-1005.

Wei W, Davis RE, Jomantiene R, Zhao Y. (2008). Ancient, recurrent phage attacks and recombination shaped dynamic sequence-variable mosaics at the root of *Phytoplasma* genome evolution. *Proc Natl Acad Sci USA* **105:** 11827-11832.

Wernegreen JJ. (2012). Endosymbiosis. *Curr Biol* **22:** R555-R561.

Xie JL, Vilchez I, Mateos M. (2010). *Spiroplasma* bacteria enhance survival of *Drosophila hydei* attacked by the parasitic wasp *Leptopilina heterotoma*. *PLoS One* **5.**

**Figure 1** Phylogenetic placement of Cameroonian and Japanese spore samples inside the Gigasporales tree. The fungal phylogeny was reconstructed using partial 28S rRNA gene sequences. The DNA sequences retrieved in this work are in bold. All the thirteen spore samples are located inside the Gigasporaceae clade, close to *Gigaspora margarita*. Supported values are from Bayesian/maximum likelihood/maximum parsimony analyses. The Bayesian and maximum likelihood analyses were performed with GTR+G nucleotide substitution model. Dashes instead numbers imply that the topology was not supported in the respective analysis.

**Figure 2** Phylogenetic placement of representative *Candidatus* Glomeribacter gigasporarum partial 16S rRNA gene sequences retrieved from spores of AMF. The DNA sequences retrieved in this work are in bold. The tree encompasses several *Ca*Gg groups. Sequences from *G. margarita*

sample CM47 and CM50 cluster in a group sister to the one (with thickened branches) including *Ca*Gg from *G. margarita* BEG34 isolate (highlighted in gray) and from the Cameroonian CM21 and CM23 samples. The 16S rRNA gene sequences from the Japanese sample MAFF520054 are located in a different and more basal position inside the tree, together with other *Ca*Gg sequences retrieved from worldwide *G. margarita* isolates. The number of sequences included in each OTU is in brackets. Cameroonian isolates showed 97-100% sequence similarity with Gigasporales isolates (*i.e. Gigaspora decipiens, G. gigantea, G. margarita*, including the isolate BEG34, *G. rosea, Racocetra castanea* and *R. verrucosa)* which are located in the upper part of the tree. By contrast, *Ca*Gg sequence similarity, in particular of the samples CM47 and CM50, decreased to 96% relative to *Ca*Gg sequences retrieved from other worldwide isolates of *Cetraspora pellucida* and *G. margarita*, including the *G. margarita* isolate MAFF520054. Supported values are from maximum likelihood/Bayesian/maximum parsimony analyses. The maximum likelihood and Bayesian analyses were performed with GTR+G and TIM3+G nucleotide substitution models, respectively. Dashes instead numbers imply that the topology was not supported in the respective analysis.

**Figure 3** Phylogenetic placement of representative *Mollicutes*-related endobacteria partial 16S rRNA gene sequences retrieved from AM spores within the *Mollicutes* clade. The DNA sequences retrieved in this work are in bold. The tree encompasses at least two main and well supported groups (*M*re group A and B) which also include sequences retrieved in previous experiments from AM spore collection (Naumann *et al*., 2010) and AMF

liverworts-associated (Desirò *et al*., 2013). The number of sequences included in each OTU is in brackets. Supported values are from Bayesian/maximum likelihood/maximum parsimony analyses. The Bayesian and maximum analyses were performed with GTR+G nucleotide substitution model. Dashes instead numbers imply that the topology was not supported in the respective analysis.

**Figure 4** Electron microscopy of *Gigaspora margarita* sample CM23. (A) The two bacterial types, *Ca*Gg (arrow) and *M*re (arrowhead) are present in the same district of the sporal fungal cytoplasm (fc). The rod-shaped type is constantly located inside a vacuole-like organelle (v). The vacuole reveals an electron dense matrix (m), identified as of protein origin. (B) Sometimes *Ca*Gg (here cut in a transversal section) is more closely surrounded by the membrane of fungal origin (arrow). (C) The *M*re is directly embedded in the fungal cytoplasm. Scale bars: (A) 1,5 μm; (B) 0,26 μm; (C) 0,17 μm.

**Figure 5** FISH on a crushed spore of *Gigaspora margarita* sample CM21. (A) Bright-field image of the fungal cytoplasm (fc) trapped in a drop of agarose is shown. (B) Triple labelling of the endobacteria with the *M*re-specific probe BLOsADf2 (red), the *Ca*Gg-specific probe CaGcADf1 (blue) and the eubacterial-probe EUB338 (green); bacteria are seen as coccoid or rod-shaped fluorescent spots (arrowheads); in this image, where red and green or blue and green channels are overlaid, bacteria are visualized as fluorescent orange or light blue spots inside the brown cytoplasm. The corresponding red, blue and green channels are shown in C, D and E. The insets show the

magnification of some *M*re and *Ca*Gg cells surrounded by the fungal cytoplasm. Scale bars: 12 µm, 3 µm in the insets.

**Figure 6** FISH on a crushed spore of *Gigaspora margarita* sample CM23. (A) Bright-field image of the fungal cytoplasm (fc) trapped in a drop of agarose is shown. (B) Triple labelling of the endobacteria with the *M*re-specific probe BLOsADf2 (red), the *Ca*Gg-specific probe CaGcADf1 (blue) and the *Buchnera*-specific probe ApisP2a (green) used as negative control; bacteria are seen as coccoid or rod-shaped fluorescent spots (arrowheads). The corresponding red and blue channels are shown in C and D. (E) No presence of non-specific fluorescent signal is detected. The insets show the magnification of some *M*re and *Ca*Gg cells surrounded by the fungal cytoplasm. Scale bars: 8 µm, 3 µm in the insets.

**Figure S1** Serial dilutions of the standard plasmids were used in individual real-time qPCR to generate standard curves for the *Ca*Gg and *M*re 16S rRNA genes. The $R^2$ values, efficiencies and slopes are shown for each reaction.

**Figure S2** Melting curve analysis of the fragments obtained by real-time q-PCR in the following conditions: (A) *Gigaspora margarita* CM23 spore DNA amplified with the *Ca*GgAD primer pair; (B) Individual plasmids carrying the three variants of the *M*re 16S rDNA amplified with the CMsAD primer pair; (C) *G. margarita* CM23 spore DNA amplified with the *M*re specific primer pair. The amplification of the sporal DNA with *Ca*Gg primers originated a unique specific melting peak. As far as the *Mollicutes*-specific amplification is

concerned, the melting analysis showed that all the three variants of the *M*re 16S rDNA could be amplified with a single primer pair, and that such different amplicons can be simultaneously obtained and discriminated when the sporal DNA is used as a template.

**Figure S3** Phylogenetic placement of Cameroonian and Japanese spore samples inside the Gigasporales tree. The fungal phylogeny was reconstructed using partial ITS sequences. The DNA sequences retrieved in this work are in bold. All the thirteen spore samples are located inside the Gigasporaceae clade, close to *Gigaspora margarita*. Supported values are from Bayesian/maximum likelihood/maximum parsimony analyses. The partitioned Bayesian analysis was performed with TVM+G, K80+G, and TPM2uf+G nucleotide substitution models for ITS1, 5.8S and ITS2 regions, respectively. The maximum likelihood analysis was performed with GTR+G nucleotide substitution model. Dashes instead numbers imply that the topology was not supported in the respective analysis.

**Figure S4** Segment-specific ML topologies reconstructed for two incongruent segments of the *M*re 16S rRNA gene sequence alignment. The breakpoint at position 479 was identified by the GARD method and is supported by the Kishino-Hasegawa test ($P < 0.001$). The trees were mid-point rooted. Values above branches represent ML bootstrap support over 70% (1,000 replicates). (A) Topology for the segment of nucleotide positions 1-479. (B) Topology for the segment of nucleotide positions 480-1109.